MOLECULAR AND CELLULAR BIOLOGY, Dec. 1989, p. 5726–5732
0270-7306/89/125726-07$02.00/0

Vol. 9, No. 12

Sequence and Structure of the *Drosophila melanogaster* Ovarian Tumor Gene and Generation of an Antibody Specific for the Ovarian Tumor Protein

WAYNE R. STEINHAUER, ROSEMARY C. WALSH, AND LAURA J. KALFAYAN*

Department of Biochemistry, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Received 8 May 1989/Accepted 31 August 1989

Sequencing cDNA and genomic DNA from the ovarian tumor gene revealed a gene with seven introns spanning 4.5 kilobases. The proline-rich, hydrophilic otu protein is novel. An antibody prepared to a β-gal-otu fusion protein recognized a 110-kilodalton ovarian protein which was altered in the ovaries of *otu* gene mutants.

The ovarian tumor (*otu*) gene product is required throughout oogenesis for the development of the female germ line (23). The 20 recessive female-sterile alleles of *otu* are classified into three categories according to the severity of their phenotypes. Ovaries from quiescent homozygotes exhibit little or no mitotic proliferation of the germ cells. Ovaries from oncogenic homozygotes undergo uncontrolled germ cell proliferation with failure of these cells to differentiate. The homozygous differentiated ovaries contain partially to fully differentiated nurse cells, oocytes, or both (21, 22, 32, 41).

While the oncogenic and differentiated alleles fail to complement the severe quiescent alleles to fertility, some heteroallelic combinations are fertile. The best example is the oncogenic/differentiated combination, otu^{11}/otu^{14}, which is fully fertile (40), suggesting that there could be more than one *otu* gene product (41) or that the product associates with itself (40) or with other molecules.

Recently, we showed that the *otu* gene hybridizes to a moderately abundant ovarian transcript of 3.2 kilobases (kb) (32). Minor ovarian RNAs of 3.8 and 4.0 kb were also detected in ovaries at a much lower abundance, and a different set of transcripts hybridizing to the *otu* gene were found in testis RNA and are also at a low abundance (about 2% of the 3.2-kb ovarian transcript; 32).

As a first step towards understanding the biochemical function of the *otu* gene product during oogenesis, we have sequenced the *otu* gene and a cDNA containing the entire coding sequence of the protein.

To isolate an *otu* cDNA clone for sequencing, an ovarian cDNA library was prepared by using poly(A)$^+$ RNA isolated from hand-dissected ovaries of Canton S flies (1, 32) which represented all stages of oogenesis. The cDNA library was prepared (15) by oligo(dT) priming, *Eco*RI linkers (sequence CCGAATTCGG) were added, and 2×10^6 recombinants were packaged in the expression vector lambda gt11 (19). The unamplified library (250,000 plaques) was screened (4) with radiolabeled probes (35). In the first screen, a probe generated to the 3.2-kb *Eco*RI fragment was used, since it hybridizes strongly to *otu* RNA (32; Fig. 1). The second and third screens were carried out by using the upstream 1.0-kb *Eco*RI fragment (32; Fig. 1) in order to select those clones with more 5′ sequences. Since the 1.0-kb *Eco*RI fragment hybridizes less efficiently to the 3.2-kb *otu* RNA on Northern (RNA) blots (32), we reasoned that the transcription unit may only extend a short way into this fragment. We identified 50 positives (0.02%), rescreened 25 of them with the 1.0-kb *Eco*RI fragment, and recovered two clones.

The cDNAs and genomic DNAs (identified previously [32] and from a Canton S library [28]) were subcloned into M13mp8 and M13mp9 (30) for sequencing by the dideoxy-chain termination method (38). The genomic restriction fragments were prepared by digestion of gel-purified *Eco*RI fragments with the following restriction enzymes: *Hae*III, *Hin*dIII, *Pst*I, *Pvu*II, and *Sau*3A (New England BioLabs, Inc., Beverly, Mass., and Promega Biotec, Madison, Wis.). DNA was sequenced by using *E. coli* DNA polymerase I (Klenow fragment; New England BioLabs) or Sequenase (United States Biochemical Corp., Cleveland, Ohio). Inserts of more than 300 bases were sequenced by using custom-synthesized oligonucleotide primers (17-mers) or by creating nested deletions (11). The genomic and cDNA sequences were assembled and analyzed with the University of Wisconsin Genetics Computer Group software.

The sequencing strategy is shown in Fig. 1A. All cDNA sequences were determined on both strands, while parts of the intron sequences were sequenced only on one strand as indicated in Fig. 1A. A few inconsistencies between the cDNA and the genome were found and could be due to polymorphisms within the Canton S flies or to errors made by the reverse transcriptase. In these cases, other overlapping cDNAs were sequenced in the questionable region. One cDNA was always found to match the genomic sequence, and it is the genomic sequence that is given in Fig. 2.

The largest *otu* cDNA (designated 3-2) is 3,108 nucleotides long and spans approximately 4.5 kb of genomic DNA. We identified eight exons which were separated by two moderately sized introns (535 and 583 base pairs) and five small introns (53 to 68 base pairs; Fig. 1B and 2). All introns were bounded by GT and AG sequences in the genome.

The cDNA is 3,108 nucleotides long, which, accounting for a poly(A) tail of average length, is in good agreement with the estimated 3.2-kb size of the mRNA (32). Since the 3.2-kb transcript is by far the most abundant ovarian transcript, it is likely that cDNA 3-2 is a close-to-full-length representative of the 3.2-kb mRNA. However, the start site of transcription is likely to be upstream of this site for the following reasons:

* Corresponding author.

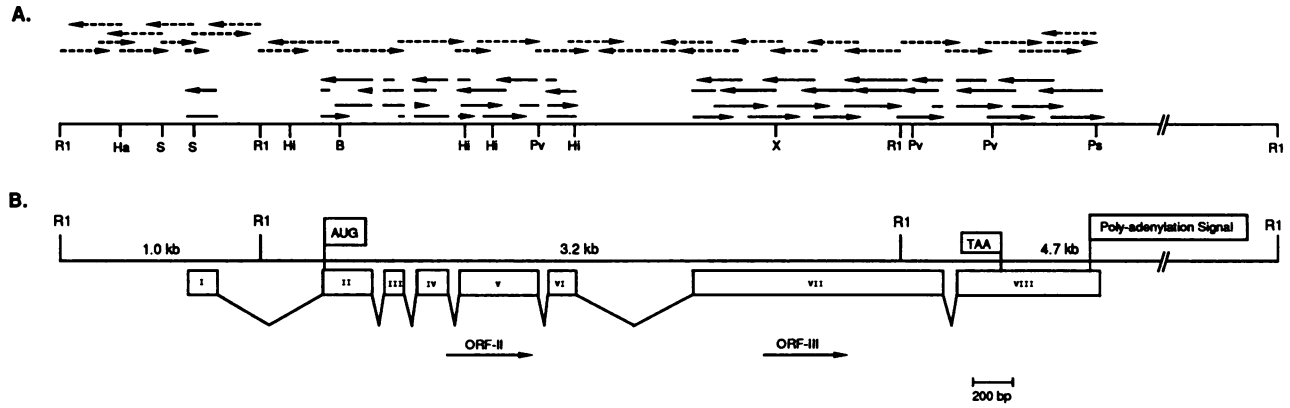

FIG. 1. *otu* gene structure. (A) Restriction map of the *otu* locus with the sequencing strategy shown above. Abbreviations: B, *Bam*HI; R1, *Eco*RI; Ha, *Hae*III; Hi, *Hin*dIII; Ps, *Pst*I; Pv, *Pvu*II; S, *Sau*3A; X, *Xho*I. The *Sau*3A and *Hae*III sites are shown only for the 1.0-kb *Eco*RI fragment at the left of the map. The 1.0- and 3.2-kb *Eco*RI fragment sizes are based on sequence data and were previously referred to as the 0.89- and 2.9-kb *Eco*RI fragments, respectively (32). cDNA sequences (solid arrows) were determined on both strands, while the overlapping genomic sequences (dashed arrows) were sequenced on both strands (the 1.0-kb *Eco*RI fragment) or on one strand (most intron sequences). (B) The genomic *Eco*RI restriction map from panel A redrawn to show the placement of exons (open boxes) and introns (crooked lines) connecting them. The exons are labeled I to VIII. Translation initiates from the AUG codon boxed above exon II and terminates at the boxed TAA above exon VIII. The positions of the polyadenylation signal AAUAAA and of two additional open reading frames (ORF-II and ORF-III) are indicated below. The arrows indicate that the direction of translation from these ORFs would be the same as that of the main ORF.

(i) the cDNA begins with a C instead of an A (or G) at a site that shares no common features with the consensus for *Drosophila melanogaster* transcriptional start sites (18); (ii) both S1 nuclease and primer extension analysis reveal a complex protection or extension pattern that suggests that the mRNA is slightly larger than the cDNA and that it possibly begins at more than one site (data not shown). The sequence GCCCAATTT, which agrees well with the consensus CAAT box sequence GGC/TCAATCT, begins 83 bases upstream from the start of the cDNA and is underlined in Fig. 2.

The presence of a poly(A) tract at the 3′ end of cDNA 3-2 and a polyadenylation signal sequence AAUAAA terminating 34 bases upstream of the poly(A) tract (Fig. 2) confirmed the location of the 3′ end of the *otu* message. The 3′ end of the *otu* message is approximately 1.4 kb from the 3′ end of the convergently transcribed *s38* chorion gene (39).

The cDNA sequence is similar to that of a cDNA isolated independently from a different library and reported recently by Champe and Laird (10). The 5′ ends are identical, and the 3′ end of our cDNA extends for 18 more nucleotides at the 3′ end. Five differences were found of the 3,076 nucleotides compared. Three were in the 3′ untranslated region of the gene, and only one would change an amino acid (our reported lysine at amino acid position 100 to a glutamine) by substituting a C for the first A of the codon.

In *D. melanogaster*, the consensus for the internal splice sequence proposed to bind to U2 mRNA is Py T Pu A Py and is located between 18 and 40 bases 5′ to the acceptor site (20). Only intron VI conforms to this consensus (Table 1). Introns I through IV and VII all contain this sequence, but it is farther upstream than normal (ranging from 45 to 68 bases upstream from the acceptor; Table 1). In introns II through IV and VII, this places the internal sequence within 10 bases of the donor consensus sequence (Fig. 2). Intron V contained two TCAAA variants of the internal consensus sequence 30 and 37 bases 5′ to the acceptor junction. This variant was also found 41 bases 5′ of the splice junction of intron VII.

A sequence of 10 A's followed by TGAAAT in intron I and by GAAAT in intron II was found 34 and 44 bases upstream from the acceptor junction (Table 1). Pyrimidine-rich stretches (31) preceded all splice acceptor sites ranging from three pyrimidines in intron VI to 10 pyrimidines in intron V (Table 1). All the 5′ donor sites showed identity in either 6 or 7 of the 9 bases making up this consensus sequence (31; Table 1). The previously reported absence of A-G's in the 19 nucleotides preceding each 3′ splice junction (20) held for all but intron I, where an A-G was found 16 nucleotides from the acceptor junction.

The variety of minor RNA species seen in ovaries and different tissues and developmental stages may be the result of differential splicing of the mRNA. For example, wild-type testes do not exhibit the predominant 3.2-kb transcript of the ovaries but instead make four minor transcripts, three larger than 3.2 kb and one smaller (32). Deletion analysis suggests that males do not require the *otu* gene product for fertility (A. Comer, unpublished observations). However, it is possible that these are alternatively spliced RNAs incapable of encoding otu protein. Differential splicing might be a mechanism of avoiding a protein that is detrimental for spermatogenesis or that would signal the female germ line developmental pathway. Other instances of sex-specific splices occur in the doublesex, Sex lethal, and *tra* genes, which are involved in somatic sex determination (2, 3, 7, 29). The higher-molecular-weight ovarian and early pupal *otu* RNAs could represent unspliced precursors, minor species with alternative start sites, developmentally significant RNAs, or any combination of these.

The cDNA 3-2 contains a long open reading frame (ORF) of 2,433 bases. Translational start and stop codons are at nucleotides 1333 and 4656, respectively, in Fig. 2. We identified two other moderately long ORFs in cDNA 3-2 (Fig. 1B). ORF-II is open from positions 1922 to 2386. It begins in exon IV and, if the RNA is spliced as in the cDNA, is open for 136 amino acids. If the RNA is not spliced at intron IV, the ORF would read through it and be open for 155 amino acids. ORF-III, in exon VII, extends from nucle-

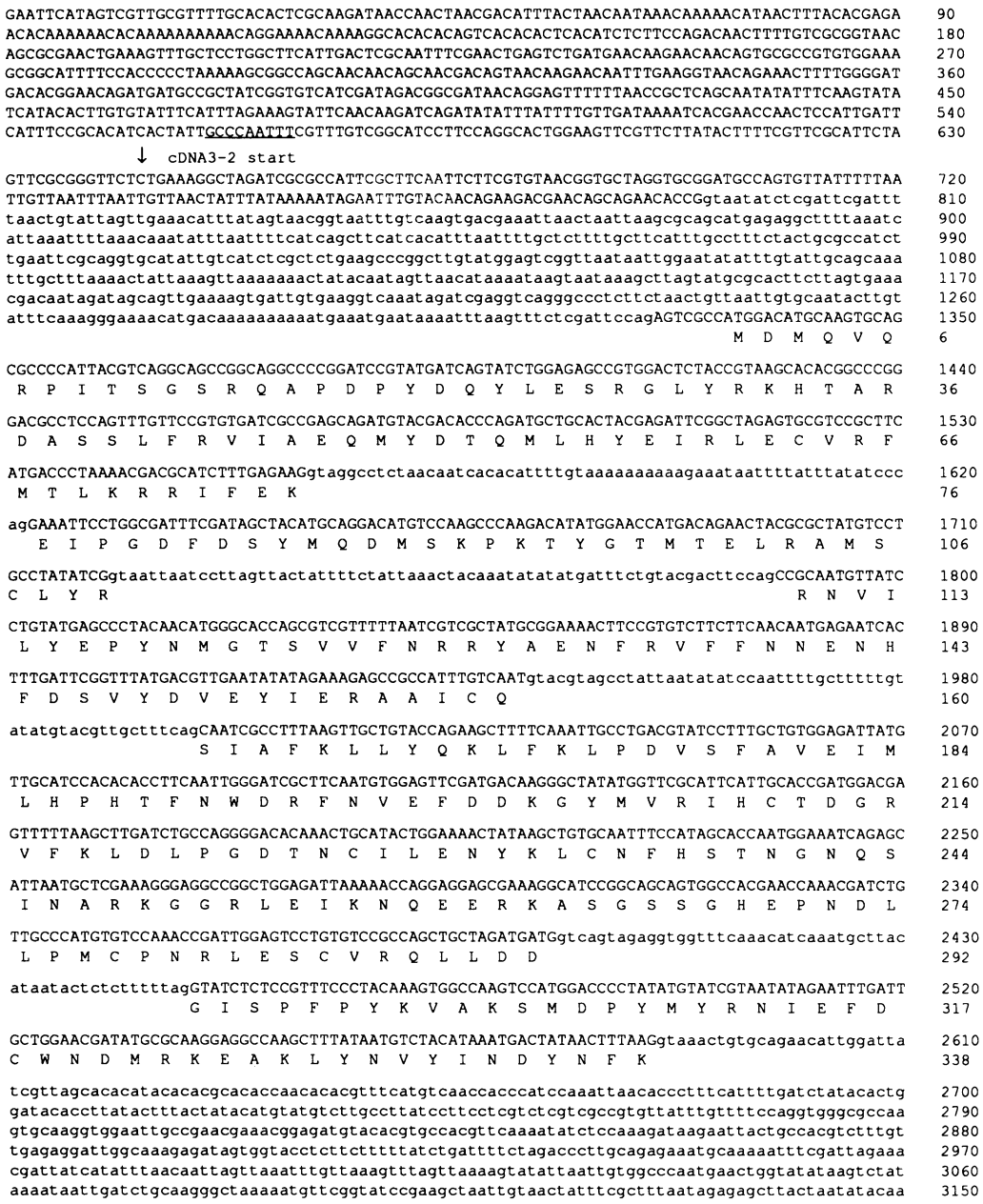

FIG. 2. DNA sequence and predicted amino acid sequence of the *otu* gene. The genomic sequence is shown along with the sequence of cDNA3-2. Intron sequences are indicated by lowercase letters. A possible CAAT box is underlined upstream of the start of cDNA 3-2 (start indicated by downward arrow). The *otu* message has a 5′ untranslated leader of at least 154 bases and begins translation at the second AUG (the first one is at position 702). A polyadenylation signal is located at nucleotides 5122 to 5127 (underlined) and a poly(A) tract is seen in the cDNA sequence. Genomic sequence was not obtained beyond the *Pst*I site (immediately preceding the boxed region), and the boxed nucleotides are from cDNA 3-2. The termination codon is marked with an asterisk.

otides 3483 to 3881 with the potential to encode 133 amino acids. Whether these ORFs are functional is unknown.

Translation appears to begin at the second start codon, at position 1333, because it begins the large open reading frame, while translation from the first AUG could only produce a 15-amino-acid peptide. Only the second AUG is surrounded by a sequence (TCGCCAUG) that resembles the consensus for eucaryotic (24) and *D. melanogaster* (9) translational initiation sites [CC(A/G)CCAUG(G) and C/AAA(C/A)AUG, respectively].

The protein predicted from cDNA 3-2 has 811 amino acids and a calculated molecular weight of 92.6 kilodaltons (Fig. 2). It is hydrophilic and has a theoretical pI of 7.2. The most striking feature of this molecule is its high proline content (12%). The prolines are not evenly distributed but are concentrated in the last two exons (VII and VIII; Table 2), which account for more than half of the protein.

The National Biomedical Research Foundation protein data bank was searched for sequence similarity to the otu protein by using the program Wordsearch (46), and no strong similarities were found by using a word size of two amino acids and an integral width of 3 for the search. We attempted

```
3151  acatatctgttggcttagGTCCCGTATGAATCGCTCCATCCCCTGCCGCCAGATGAGTACCGCCCATGGTCGTTGCCATTCCGCTATCAT   3240
 339                      V  P  Y  E  S  L  H  P  L  P  P  D  E  Y  R  P  W  S  L  P  F  R  Y  H     362

3241  CGCCAGATGCCTCGCTTGCCGTTGCCCAAGTATGCCGGTAAGGCCAACAAGTCTTCCAAATGGAAGAAGAACAAGCTGTTCGAAATGGAC   3330
 363   R  Q  M  P  R  L  P  L  P  K  Y  A  G  K  A  N  K  S  S  K  W  K  K  N  K  L  F  E  M  D     392

3331  CAGTATTTTGAGCACAGCAAGTGTGATTTGATGCCCTACATGCCCGTGGACAATTGCTATCAGGGTGTGCACATTCAGGACGATGAGCAG   3420
 393   Q  Y  F  E  H  S  K  C  D  L  M  P  Y  M  P  V  D  N  C  Y  Q  G  V  H  I  Q  D  D  E  Q     422

3421  CGGGATCATAATGATCCTGAACAAAATGACCAGAACCCGACTACGGAGCAGCGGGATCGTGAAGAACCGCAGGCACAGAAGCAACACCAG   3510
 423   R  D  H  N  D  P  E  Q  N  D  Q  N  P  T  T  E  Q  R  D  R  E  E  P  Q  A  Q  K  Q  H  Q     452

3511  CGCACGAAGGCATCAAGGGTTCAGCCGCAGAACTCGAGTTCCAGCCAAAACCAGGAGGTTTCGGGTTCGGCTGCCCCGCCACCCACTCAG   3600
 453   R  T  K  A  S  R  V  Q  P  Q  N  S  S  S  S  Q  N  Q  E  V  S  G  S  A  A  P  P  P  T  Q     482

3601  TATATGAATTACGTGCCAATGATACCGAGTCGTCCTGGGCATTTACCGCCACCTTGGCCTGCATCTCCGATGGCTATTGCCGAGGAGTTT   3690
 483   Y  M  N  Y  V  P  M  I  P  S  R  P  G  H  L  P  P  P  W  P  A  S  P  M  A  I  A  E  E  F     512

3691  CCGTTCCCCATTTCAGGAACCCCGCATCCACCGCCAACCGAAGGTTGTGTATACATGCCATTCGGTGGTTATGGTCCACCACCACCGGGA   3780
 513   P  F  P  I  S  G  T  P  H  P  P  P  T  E  G  C  V  Y  M  P  F  G  G  Y  G  P  P  P  P  G     542

3781  GCTGTTGCTTTATCGGGACCGCATCCATTTATGCCGCTTCCTTCTCCACCGCTAAATGTTACCGGAATTGGCGAGCCACGTCGTTCTCTA   3870
 543   A  V  A  L  S  G  P  H  P  F  M  P  L  P  S  P  P  L  N  V  T  G  I  G  E  P  R  R  S  L     572

3871  CACCCAAACGGTGAAGATTTGCCCGTGGATATGGTGACTTTGAGATACTTCTACAACATGGGCGTGGATTTGCATTGGCGCATGTCGCAC   3960
 573   H  P  N  G  E  D  L  P  V  D  M  V  T  L  R  Y  F  Y  N  M  G  V  D  L  H  W  R  M  S  H     602

3961  CACACGCCGCCTGATGAACTAGGAATGTTTGGATACCATCAGCAGAACAACACTGATCAACAGGCAGGACGGACTGTAGTCATTGGCGCC   4050
 603   H  T  P  P  D  E  L  G  M  F  G  Y  H  Q  Q  N  N  T  D  Q  Q  A  G  R  T  V  V  I  G  A     632

4051  ACAGAGGACAATTTGACTGCCGTGGAGTCAACACCACCACCTTCGCCAGAGGTGGCAAATGCCACAGAGCAGTCACCGCTTGAGAAAAGT   4140
 633   T  E  D  N  L  T  A  V  E  S  T  P  P  P  S  P  E  V  A  N  A  T  E  Q  S  P  L  E  K  S     662

4141  GCCTACGCCAAGCGCAATTTGAATTCGGTTAAGGTGCGCGGCAAACGTCCGGAGCAGCTGCAAGATATTAAGGATTCGCTGGGGCCAGCG   4230
 663   A  Y  A  K  R  N  L  N  S  V  K  V  R  G  K  R  P  E  Q  L  Q  D  I  K  D  S  L  G  P  A     692

4231  GCATTTTTGCCCACTCCAACGCCATCGCCAAGCTCGAATGGCAGTCAGTTTAGTTTCTATACTACTCCATCGCCGCATCATCACCTGATA   4320
 693   A  F  L  P  T  P  T  P  S  P  S  S  N  G  S  Q  F  S  F  Y  T  T  P  S  P  H  H  H  L  I     722

4321  ACACCGCCGAGGTTGCTCCAACCGCCGCCACCGCCACCGATATTCTACCACAAGGCGGGACCACCACAGCTAGGGGGAGCAGCTCAAGGA   4410
 723   T  P  P  R  L  L  Q  P  P  P  P  P  P  I  F  Y  H  K  A  G  P  P  Q  L  G  G  A  A  Q  G     752

4411  CAGgtaggagtgatacatgcactaacaaattcaaaatattctataggcaatcgacactcgaccatttttagACTCCCTACGCCTGGGGCA   4500
 753   Q                                                                    T  P  Y  A  W  G     760

4501  TGCCAGCTCCGGTGGTGTCCCCCTATGAGGTGATCAACAACTATAACATGGACCCGTCGGCTCAGCCACAACAACAGCAGCCAGCCCCCT   4590
 761   M  P  A  P  V  V  S  P  Y  E  V  I  N  N  Y  N  M  D  P  S  A  Q  P  Q  Q  Q  Q  P  A  P     790

4591  TGCAACCAGCTCCCTTATCTGTCCAATCTCAGCCGGCAGCTGTCTATGCTGCAACGCGTCATCACTAAACAAAGAAAGAGAAAAAAAAGG   4680
 791   L  Q  P  A  P  L  S  V  Q  S  Q  P  A  A  V  Y  A  A  T  R  H  H  *                       811

4681  GAGCGGGGGCAAAAAACAGATCACTTGAAAGAGAGAGGCCATACAGATCGAAGGCACTACATTCCATTGCAATTAACGGCTTTTAAAATT   4770
4771  TAATCTCACTTTTAAATTTGTAGTTAACTTTTTATAGGCCATAAGCGTTGGCGCTCTATCATAAACCATTCAGCTTCTGTACAACAATCG   4860
4861  ATTGCATAACCTAACGCAAATGTCAACCCAACTTCATTTTAAAAATGTAATTTAACGTAATTTTATGCGAATTTTTTTAAAGTTAGCCGT   4950
4951  CACGAAATCAAAGAACCACCTATTTATATGATTTATTTAAAACCCTTCTAACCAAAAATATCTACATACTATCTACTATATATATACATA   5040
5041  TATATATATATATATATTTATGTGCTCGCTGTTCGGCTAGAGACTCACCTATGTAAAGTGTACCATCAAAAATTAACCATAAATAAAACA   5130
5131  AGATTCAACTGCAG CGCAAGAGACAAAATGTAAAAAAAAAAAAAAA
```

FIG. 2—*Continued.*

TABLE 1. Consensus sequences in introns I through VII

Intron	Donor consensus $\frac{C}{A}$AG GT$\frac{G}{A}$AGT	Internal consensus signal $\frac{C}{T}$T$\frac{A}{G}$A$\frac{C}{T}$	Acceptor consensus Pyr stretch-AG	Conserved A-rich sequence
I	CCG GTATAT	−90 CTAAC −83 TTAAT −68 TTGAT	TTCCAG	−44 AAAAAAAAAATGAAAT
II	AAG GTAGGC	−54 CTAAC	TCCCAG	−34 AAAAAAAAAA GAAAT
III	TCG GTAATT	−63 TTAAT	CTTCCAG	
IV	AAT GTACGT	−45 TTAAT −36 CCAAT	CTTTCAG	
V	ATG GTCAGT	−37 TCAAA −30 TCAAA	CTCTCTTTTTAG	
VI	AAG GTAAAC	−44 TTAAT −29 CTAAT	CTTAG	
VII	CAG GTAGGA	−50 CTAAC −41 TCAAA	TTTTTAG	

TABLE 2. Distribution of prolines in the translated exons of *otu*

Exon	Proline/amino acid residue ratio	Proline content (%)	Amino acid positions
II	5/76	6.5	1–76
III	2/33	6.06	77–109
IV	1/51	2	110–160
V	4/132	3.05	161–292
VI	3/46	6.66	293–338
VII[a]	71/415	17.15	339–753
VIII	10/58	19.6	754–811

[a] A subregion of exon VII has a proline/amino acid residue ratio of 31/111, its proline content is 27.9%, and it is found at amino acid positions 477 to 587.

to assign the otu protein to a functional class of molecule by searching for a variety of specific domain consensus sequences. The otu protein did not contain consensus sequences for ATP-binding sites (45), helicases (17), RNA-binding sites (34, 42), leucine zippers (27), or DNA-binding sites of the zinc finger (13) or GCN4 class (44). A hydropathy analysis (25) showed a region of modest hydrophobicity, but the average hydropathy index over 19 amino acids (<1.0) was considerably less than the value of 1.6 expected for a membrane-spanning domain (25).

In order to begin to analyze the biochemical function of the otu protein, we have generated an antibody to use as a reagent for localization of the normal protein and analysis of mutant proteins. A partial *otu* cDNA obtained from S. Parks (33) containing sequences that encode 421 amino acids from residues 253 to 671 was inserted into an expression vector, pWR590 (16), in frame with part of the β-galactosidase gene. The construct was transformed into *Escherichia coli* MV1189 cells, and it expressed a β-gal-otu fusion protein of approximately 120 kilodaltons which was absent in cells transformed with pWR590 alone. The protein was partially purified by insoluble aggregation (36, 47). The enriched preparation was used to immunize rabbits (36) to generate antibodies against the otu protein. Antisera was affinity purified to remove β-galactosidase antibodies (37) and subsequently to enrich for antibodies that bound to the β-gal-otu fusion protein. The purified antibody was tested for specificity to *D. melanogaster* ovarian proteins by Western blot (immunoblot) analysis (26, 43).

Ovarian proteins were prepared by homogenizing ovaries in buffer (50 mM Tris hydrochloride [pH 7.5], 3 mM EDTA, 1% Nonidet P-40, 0.1% sodium dodecyl sulfate) containing 1 mM *N*-ethylmaleimide, 100 μm leupeptin, 10 μm pepstatin, 3 U of trypsin inhibitor per ml of aprotinin, and 100 μg of phenylmethylsulfonyl fluoride. After centrifugation at 13,000 × *g* for 3 min, the supernatants were denatured, electrophoresed, and analyzed on Western blots (5, 26, 43).

An ovarian protein of approximately 110 kilodaltons was detected with the anti-otu antibody (Fig. 3). This protein was greatly reduced in the ovaries of the DIF allele otu^{P4} and absent in the ovaries of the DIF allele otu^{14} and was not detected by preimmune sera (Fig. 3B). A new protein of approximately 88-kilodaltons, detected in otu^{14} but absent in wild-type or otu^{P4} flies, may represent a truncated mutant form of the otu protein. In addition, another protein of slightly higher molecular weight was present in some of the lanes. The diffuse band of approximately 45 kilodaltons may be a breakdown product of the 110-kilodalton protein or may represent nonspecific binding of the antibody to the highly abundant vitellogenin proteins of the ovaries. The 26-kilodalton band seen in all panels is due to cross-reactivity of the secondary antibody to an ovarian protein (Fig. 3C). The size of the protein identified (110 kilodaltons) is larger than the predicted 92.6-kilodalton otu protein. It is possible that this protein is posttranslationally modified or that the apparent higher molecular weight is the result of the high proline content as has been reported for the fushi tarazu, bicoid, and Krüppel proteins of *D. melanogaster* (8, 12, 14).

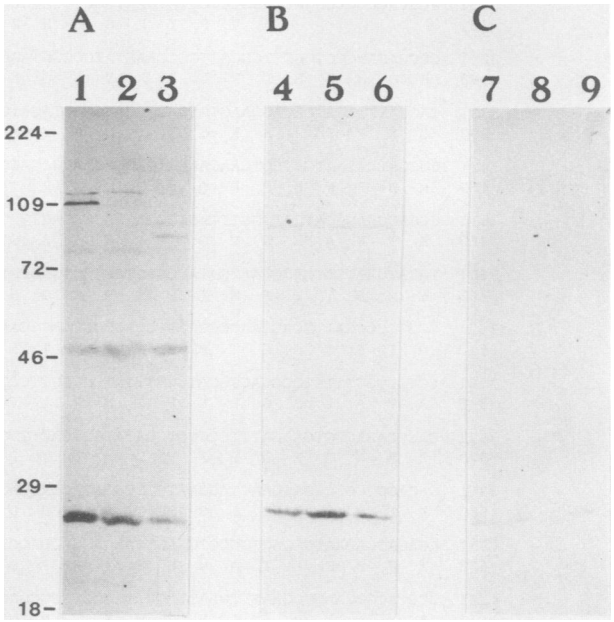

FIG. 3. Specificity of affinity-purified anti-*otu* antisera. Ovarian proteins were extracted from wild-type (Canton S), otu^{P4}, and otu^{14} flies. Protein concentrations were determined by the Bio-Rad protein assay, which is based on the Bradford assay (6). Each lane was loaded with 50 μg of protein, electrophoresed on 10% SDS-sodium dodecyl sulfate-polyacrylamide gels, and transferred to nitrocellulose. Blots were incubated with affinity-purified anti-*otu* antibody (1:35,000 dilution) (A), preimmune serum (1:35,000 dilution) (B), or secondary antibody only (C). Lanes 1, 4, and 7 have wild-type proteins. Lanes 2, 5, and 8 have otu^{P4} proteins, and lanes 3, 6, and 9 have otu^{14} protein. Proteins reacting with the antibody were detected by using alkaline phosphatase-conjugated goat anti-rabbit antibody (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) and the color reagents 5-bromo-4-chloro-3-indolyl-phosphate-*p*-toluidine salt) and *p*-Nitro Blue Tetrazolium chloride). Molecular sizes (in kilodaltons) are given at the right and are based on prestained protein molecular size markers (BioRad Laboratories, Richmond, Calif.).

We thank David Joseph and Patrick Sullivan for preparing the ovarian cDNA library and gratefully acknowledge the support of the Laboratory of Reproductive Biology's Core Recombinant DNA Facility (supported by Public Health Service grant 5-P30-HD-1898 from the National Institutes of Health to Frank French) for the library construction. We are grateful to Mark Champe and Charles Laird for sharing their *otu* sequence data prior to publication. We thank Gustavo Maroni, Gwen Sancar, Ron Swanstrom, and Bob King for constructive criticism of the manuscript. We thank Dana Fowlkes in the Department of Pathology, University of North Carolina at Chapel Hill, for synthesizing the oligonucleotides used in this work.

This work was funded by Public Health Service grant GM36801 (to L.J.K.) from the National Institutes of Health and also by grant NP-657 (to L.J.K.) from the American Cancer Society.

LITERATURE CITED

1. **Aviv, J., and P. Leder.** 1972. Purification of biologically active globin mRNA by column chromatography on oligothymidillic acid cellulose. Proc. Natl. Acad. Sci. USA **69**:1408–1412.
2. **Baker, B. S., and M. F. Wolfner.** 1988. A molecular analysis of *doublesex*, a bifunctional gene that controls both male and female sexual differentiation in *Drosophila melanogaster*. Genes & Dev. **2**:477–489.
3. **Bell, L. R., E. M. Maine, P. Schedl, and T. W. Cline.** 1988. *Sex-lethal*, a Drosophila sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. Cell **55**:1037–1046.
4. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques *in situ*. Science **196**:180–182.
5. **Bjerrum, O. J., and C. Schafer-Nielsen.** 1986. Analytical electrophoresis, p. 315. Verlag Chemie, Weinheim, Federal Republic of Germany.
6. **Bradford, M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
7. **Butler, B., V. Pirotta, I. Irminger-Finger, and R. Nothiger.** 1986. The sex-determining gene *tra* of *Drosophila melanogaster*: molecular cloning and transformation studies. EMBO J. **5**: 3607–3613.
8. **Carroll, S. B., and M. P. Scott.** 1985. Localization of the *fushi tarazu* protein during Drosophila embryogenesis. Cell **43**:47–57.
9. **Cavener, D. R.** 1987. Comparison of the consensus sequence flanking translational start sites in Drosophila and vertebrates. Nucleic Acids Res. **15**:1353–1361.
10. **Champe, M. A., and C. D. Laird.** 1989. Nucleotide sequence of a cDNA from the putative *ovarian tumor* locus of *Drosophila melanogaster*. Nucleic Acids Res. **17**:3304.
11. **Dale, R. M. K., B. A. McClure, and J. P. Houchins.** 1985. A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. Plasmid **13**:31–40.
12. **Driever, W., and C. Nusslein-Volhard.** 1988. A gradient of bicoid protein in Drosophila embryos. Cell **54**:83–93.
13. **Evans, R. M., and S. M. Hollenberg.** 1988. Zinc fingers: gilt by association. Cell **52**:1–3.
14. **Gaul, U., E. Seifert, R. Schuh, and H. Jäckle.** 1987. Analysis of Krüppel protein distribution during early Drosophila development reveals posttranscriptional regulation. Cell **50**:639–647.
15. **Gubler, U., and B. J. Hoffman.** 1983. A simple and very efficient method for generating cDNA libraries. Gene **25**:263–269.
16. **Guo, L., P. P. Stepien, J. Y. Tso, R. Broussequ, S. Narang, D. Y. Thomas, and R. Wu.** 1984. Synthesis of human insulin gene. VIII. Construction of expression vector for fused proinsulin production in *Escherichia coli*. Gene **29**:251–254.
17. **Hodgman, T. C.** 1988. A new superfamily of replicative proteins. Nature (London) **333**:578.
18. **Hultmark, D., R. Klemenz, and W. J. Gehring.** 1986. Translational and transcriptional control elements in the untranslated leader of the heat-shock gene *hsp22*. Cell **44**:429–438.
19. **Huynh, T. V., R. A. Young, and R. W. Davis.** 1985. Constructing and screening cDNA libraries in λgt10 and λgt11, p. 49–78. *In* D. Glover, (ed.), DNA cloning techniques: a practical approach. IRL Press, Oxford.
20. **Keller, E. B., and W. A. Noon.** 1985. Intron splicing: a conserved internal signal in introns of *Drosophila* pre-mRNAs. Nucleic Acids Res. **13**:4971–4981.
21. **King, R. C., J. D. Mohler, S. F. Riley, P. D. Storto, and P. D. Nicolazzo.** 1986. Complementation between alleles at the *ovarian tumor* (*otu*) locus of *Drosophila melanogaster*. Dev. Genet. **7**:1–20.
22. **King, R. C., and S. F. Riley.** 1982. Ovarian pathologies generated by various alleles of the *otu* locus in *Drosophila melanogaster*. Dev. Genet. **3**:69–89.
23. **King, R. C., and P. D. Storto.** 1988. The role of the *otu* gene in Drosophila oogenesis. BioEssays **8**:18–24.
24. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12**:857–872.
25. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
26. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
27. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240**:1759–1764.
28. **Maniatis, T., R. C. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quon, G. K. Sim, and A. Efstradiadis.** 1978. The isolation of structural genes from libraries of eucaryotic DNA. Cell **15**: 687–701.
29. **McKeown, M., J. M. Belote, and R. T. Boggs.** 1988. Ectopic expression of the female *transformer* gene product leads to female differentiation of chromosomally male *Drosophila*. Cell **53**:887–895.
30. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
31. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
32. **Mulligan, P. K., J. D. Mohler, and L. J. Kalfayan.** 1988. Molecular localization and developmental expression of the *otu* locus of *Drosophila melanogaster*. Mol. Cell. Biol. **8**:1481–1488.
33. **Parks, S., and A. Spradling.** 1987. Spatially regulated expression of chorion genes during *Drosophila* oogenesis. Genes & Dev. **1**:497–509.
34. **Prengschat, F., and B. Wold.** 1988. Isolation and characterization of a *Xenopus laevis* C protein cDNA: structure and expression of a heterogeneous nuclear ribonucleoprotein core protein. Proc. Natl. Acad. Sci. USA **85**:9669–9673.
35. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labelling of deoxyribonucleic acid to high specific activity *in vitro* by nick-translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
36. **Rio, D. C., F. A. Laski, and G. M. Rubin.** 1986. Identification and immunochemical analysis of biologically active *Drosophila* P element transposase. Cell **44**:21–32.
37. **Robbins, A., W. S. Dynan, A. Greenleaf, and R. Tjian.** 1984. Affinity-purified antibody as a probe of RNA polymerase II subunit structure. J. Mol. Appl. Genet. **2**:343–353.
38. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
39. **Spradling, A. C., D. V. deCicco, B. T. Wakimoto, J. F. Levine, L. J. Kalfayan, and L. Cooley.** 1987. Amplification of the X-linked Drosophila chorion gene cluster requires a region upstream from the *s38* chorion gene. EMBO J. **6**:1045–1053.
40. **Storto, P. D., and R. C. King.** 1987. Fertile heteroallelic combinations of mutant alleles of the *otu* locus of *Drosophila melanogaster*. Roux's Arch. Dev. Biol. **196**:210–221.
41. **Storto, P. D., and R. C. King.** 1988. Multiplicity of functions for the *otu* gene products during *Drosophila* oogenesis. Dev. Genet. **9**:91–120.
42. **Swanson, M. S., T. Y. Nakagawa, K. LeVan, and G. Dreyfuss.** 1987. Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-mRNA-binding proteins. Mol. Cell. Biol. **7**:1731–1739.
43. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
44. **Vogt, P. K., T. J. Bos, and R. F. Doolittle.** 1987. Homology between the DNA-binding domain of the GCN4 regulatory

protein of yeast and carboxy-terminal region of a protein coded for by the oncogene *jun*. Proc. Natl. Acad. Sci. USA **84:** 3316–3319.

45. **Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay.** 1982. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1:**945–951.
46. **Wilbur, W. J., and D. J. Lipman.** 1983. Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA **80:**726–730.
47. **Williams, D. C., R. M. Van Frank, W. L. Muth, and J. P. Burnett.** 1982. Cytoplasmic inclusion bodies in *Escherichia coli* producing biosynthetic human insulin proteins. Science **215:** 687–689.